

# Joint Interpolation and Approximation of Causal Systems\*

WILLIAM A. PORTER

*Department of Electrical Engineering, Louisiana State University,  
Baton Rouge, Louisiana 70803*

*Submitted by C. L. Dolph*

The modeling of physical systems inherently involves constructing a mathematical approximation from observable data and/or a priori assumptions. This study refines some recent work on causal interpolation and causal approximation as system modeling techniques. Sufficient conditions for causal interpolators to approximate continuous causal systems are established. State realizations for minimal norm causal interpolators are also established.

## 1. INTRODUCTION

The identification and/or representation of a “black-box” phenomenon from external measurements is a modeling problem of contemporary interest. The relevant literature gives ample testimony to the diversity of problems and techniques that come within the bounds of this problem class.

Automatic control theorists, for example, have been pursuing identification problems for linear dynamic systems (see survey [1]) for several years. In the nonlinear setting the representation of a black box by polynomial or multilinear models has had a recent resurgence of interest (see survey [2]). Such an approach is essential where the input–output behavior is nonlinear even for small signals. A recent article by Palm and Poggio [3] has underscored the importance of polynomial modeling, i.e., Volterra–Wiener expansions, in the biological systems domain.

The present paper utilizes a Hilbert space setting and considers two distinct problems.

The “interpolation problem” can be summarized as follows. We start with a collection of observed input–output pairs  $\{(u_i, y_i): i = 1, \dots, m\}$ . From these a map,  $\Phi$ , is sought such that

$$y_i = \Phi(u_i), \quad i = 1, \dots, m. \quad (1)$$

If the pairs are derived from experimental observation then the map  $\Phi$  is obviously a representation of the black box, valid over existing data.

\* Supported in part under AFOSR Grant 78-3500.

The interpolation problem can be augmented with various constraints, for example, that

- (i)  $\Phi$  be continuous,
  - (ii)  $\Phi$  be linear,
  - (iii)  $\Phi$  be causal,
  - (iv)  $\Phi$  be polynomial,
  - (v)  $\Phi$  be of minimum norm,
  - (vi)  $\Phi$  satisfy a priori structure.
- (2)

With appropriate choice of constraints several potential applications come in view. First there is a strong connection with pattern identification problems, i.e.,  $\{u_i\}$  are the patterns,  $\{y_i\}$  are the signatures, and  $\Phi$  is the pattern recognizer. Second, there is a connection with fault detection wherein  $\{u_i\}$  are possible fault categories and  $\{y_i\}$  are the desired fault signatures. More recently it has been determined that a form of the interpolation problem can be applied to the design of adaptive systems [4] and observers [5].

The second problem of interest is the "approximation problem." Here an input domain,  $K$ , is specified for the black-box system,  $f$ . Input-output data  $\{(u_i, y_i): i = 1, \dots, m\}$  are also in hand and the problem is to choose  $\Psi$  such that

$$\Psi \simeq f$$

on the domain  $K$ . The constraints of Eq. (2) are also useful constraints for the approximation problem. For example, with the constraints equation (2vi) being a linear dynamical structure, the approximation problem posed above is consistent with the literature touched on in [1].

While the interpolation and approximation problems have an obvious similarity it is clear that approximating systems need not interpolate and conversely interpolating systems need not approximate. The present study establishes conditions under which the dual property holds, that is, the interpolating system does also approximate.

## 2. MATHEMATICAL PRELIMINARIES

It is convenient to focus specifically on the Hilbert space  $L_2(v)$  over the finite or infinite interval,  $v$ , and equipped with the usual inner product. We shall need also the orthoprojector family  $\{p^t: t \in v\}$  given by

$$\begin{aligned} (P x)(\beta) &= x(\beta), & \beta \leq t, \\ &= 0, & \beta > t. \end{aligned} \tag{3}$$

A function  $f: L_2(v) \rightarrow L_2(v)$  is *causal* provided  $P^t f = P^t f P^t$ , all  $t \in v$ .

The set  $\{(u_i, y_i): i = 1, \dots, m\}$  is said to be *linearly well posed* if it satisfies the condition

$$\sum_{i=1}^m \alpha_i(P^t u_i) = 0 \Rightarrow \sum_{i=1}^m \alpha_i(P^t y_i) = 0, \quad \text{all } t \in v. \quad (4)$$

In [6] it is shown that linearly well-posed sets admit a causal linear map,  $\Phi$ , satisfying Eq. (1). In addition one such map is explicitly constructed. In [7] the linear solution is shown to have a state realization embodied in a family of  $m$  linear differential equations.

When the set  $\{(u_i, y_i): i = 1, \dots, m\}$  is not linearly well posed then [6] provides a causal polynomial map satisfying the input-output constraints. This polynomial map is of order  $m - 1$  (we shall clarify "order" later) and is by no means unique.

In the same setting [8, 9], consider the polynomial approximation of continuous functions on  $L_2(v)$ . In particular if  $K \subset L_2(v)$  is an arbitrary compact set and if  $f$  is a continuous causal function on  $L_2(v)$  then [8] shows the existence of a causal polynomial map,  $\hat{f}$ , on  $L_2(v)$  such that

$$\sup_{u \in K} \|f(u) - \hat{f}(u)\| < \epsilon$$

holds for arbitrary  $\epsilon > 0$ . Moreover, [9] shows that  $\hat{f}$  has a state variable realization which is linear in state behavior and polynomial in its state-to-output map.

### Vector Extensions

In addition to  $H = L_2(v)$  it is helpful to introduce several sets and spaces related to  $H$ . First the notation  $H(j)$  is defined by

$$H(j) = \{x: x \text{ is square integrable over } v^j\}, \quad j = 0, 1, \dots, n.$$

The space  $H(0)$  denotes the scalar field. For  $p, q \in H(j)$  we have

$$\langle p, q \rangle(j) = \int_{v^j} p(s_1 \cdots s_j) q(s_1 \cdots s_j) d\mu(s_1) \cdots d\mu(s_j), \quad j = 1, \dots, n. \quad (5)$$

The case  $j = 0$  is taken as scalar multiplication and  $H(1) = H$ .

Two constructions are of interest. Consider first the product space

$$\mathcal{H} = H(0) \times H(1) \times \cdots \times H(n)$$

which is equipped with the inner product

$$(x, y) = x_0 y_0 + \langle x_1, y_1 \rangle(1) + \cdots + \langle x_n, y_n \rangle(n), \quad (6)$$

where  $x_j, y_j \in H(j)$ ,  $j = 0, 1, \dots, n$ , and

$$\begin{aligned}x &= (x_0, x_1, \dots, x_n), \\y &= (y_0, y_1, \dots, y_n).\end{aligned}$$

It is obvious that  $\mathcal{H}$  is a Hilbert space.

The second construction is denoted by  $\mathbf{H}$ . For arbitrary  $x \in H$  we form the vector

$$\mathbf{x} = (1, x, x^2, \dots, x^n), \quad (7)$$

where

$$x^j(\alpha_1 \cdots \alpha_j) = x(\alpha_1) x(\alpha_2) \cdots x(\alpha_j), \quad \alpha_j \in v, i = 1, \dots, n.$$

It is clear that  $\mathbf{H} \subset \mathcal{H}$  and that this inclusion is proper. For  $\mathbf{x}, \mathbf{y} \in \mathbf{H}$  it follows easily that

$$(\mathbf{x}, \mathbf{y}) = 1 + \langle x_1, y_1 \rangle(1) + \cdots + \langle x_n, y_n \rangle(n).$$

We note also that  $\mathbf{H}$  is not a linear subspace.

Consider now an arbitrary set  $K \subset H$ . The set  $\mathbf{K} = \{\mathbf{x}: x \in K\}$  is formed using Eq. (3).

LEMMA 1. *If  $K$  is closed in  $H$  then  $\mathbf{K}$  is closed in  $\mathcal{H}$ .*

*Proof.* Let  $\{\mathbf{Z}^i\} \subset \mathbf{K}$  denote a sequence with limit  $\xi \in \mathcal{H}$ . For every  $\epsilon > 0$  there exist  $k$  such that (here  $\|\cdot\|$  is the norm on  $H$ )

$$\epsilon^2 \geq \|\mathbf{Z}^i - \xi\|^2 = \|\xi_0 - 1\|^2 + \|\xi_1(\cdot) - Z^i(\cdot)\|^2 + \|\xi_2(\cdot) - Z^i(\cdot)Z^i(\cdot)\|^2 + \cdots$$

for all  $i \geq k$ . It follows immediately that

$$\|\xi_j(\cdot, \dots, \cdot) - Z^i(\cdot), \dots, Z^i(\cdot)\| < \epsilon, \quad j = 0, 1, \dots, n.$$

In particular for  $K$  closed we have that  $\xi_1 \in K$ . Now using the equality

$$\begin{aligned}\xi_2(\alpha, \beta) - \xi_1(\alpha) \xi_1(\beta) &= \xi_1(\alpha, \beta) - Z^i(\alpha) Z^i(\beta) \\ &\quad + Z^i(\alpha)[Z^i(\beta) - \xi_1(\beta)] + \xi_1(\beta)[Z^i(\alpha) - \xi_1(\alpha)]\end{aligned}$$

and standard norm inequalities it follows that

$$\|\xi_1(\cdot, \cdot) - \xi_1(\cdot) \xi_1(\cdot)\| \leq \epsilon(1 + 2\|Z\|)$$

and hence, since  $\epsilon$  is arbitrary,

$$\xi_2(\cdot, \cdot) = \xi_1(\cdot) \xi_1(\cdot).$$

An immediate corollary of Lemma 1 is that  $\mathbf{H}$  is closed in  $\mathcal{H}$ .

Consider now  $x, y \in H$  and associated vectorized forms  $\mathbf{x}, \mathbf{y} \in \mathbf{H}$ . The identities

$$\begin{aligned}\|\mathbf{x} - \mathbf{y}\|^2 &= \|x - y\|^2 + \cdots + \|x - y\|^{2n}, \\ \|\mathbf{x}\|^2 &= 1 + \|x\|^2 + \cdots + \|x\|^{2n}\end{aligned}\tag{9}$$

follow immediately from Eqs. (6) and (7). When  $\|x - y\| \leq 1$ , respectively  $\|x\| < 1$ , the closed forms

$$\begin{aligned}\|\mathbf{x} - \mathbf{y}\|^2 &= \|x - y\|^2 \frac{1 - \|x - y\|^{2n}}{1 - \|x - y\|^2}, \\ \|\mathbf{x}\|^2 &= \frac{1 - \|x\|^{2n+2}}{1 - \|x\|^2}\end{aligned}\tag{10}$$

are also easily verified.

**LEMMA 2.** *If  $K$  is compact in  $H$  then  $\mathbf{K}$  is compact in  $\mathcal{H}$ .*

*Proof.* Using Eq. (10) it follows easily that if  $\{x_1, \dots, x_k\}$  is an  $\epsilon$ -net for  $K$ , with  $0 < \epsilon < \frac{1}{2}$ , then  $\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$  is a  $2\epsilon$ -net for  $\mathbf{K}$ .

### 3. FUNCTION APPROXIMATION

Let  $\tau: H \rightarrow \mathbf{H}$  denote the map  $\tau(x) = \mathbf{x}$  computed in Eq. (7). It is clear that  $\tau$  is 1:1, onto, and *not* linear.

For every function  $f: H \rightarrow H$  there is an associated function  $\mathbf{f}: \mathbf{H} \rightarrow H$  computed by  $\mathbf{f}\tau = f$ . When  $f$  is Lipschitz on  $H$  we use the symbol

$$|f| = \sup_{x \neq y} \frac{\|f(x) - f(y)\|}{\|x - y\|}, \quad x, y \in H,$$

to denote the Lipschitz norm. By a slight abuse of notation we use also

$$|g| = \sup_{\mathbf{x} \neq \mathbf{y}} \frac{\|g(\mathbf{x}) - g(\mathbf{y})\|}{\|\mathbf{x} - \mathbf{y}\|}, \quad \mathbf{x}, \mathbf{y} \in \mathbf{H},$$

whenever  $g: \mathbf{H} \rightarrow H$  and the indicated sup is finite.

**LEMMA 3.** *If  $f$  is Lipschitz on  $K \subset H$  then  $\mathbf{f}$  is Lipschitz from  $\tau(K) = \mathbf{K}$  into  $H$  and  $|\mathbf{f}| \leq |f|$ .*

*Proof.* The lemma follows from the identity

$$\frac{\|f(x) - f(y)\|}{\|x - y\|} = \frac{\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\|}{\|x - y\|} \geq \frac{\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\|}{\|\mathbf{x} - \mathbf{y}\|}$$

and the onto property of  $\tau$ .

Consider then  $K \subset H$ , where  $K$  is compact and  $\{x_1, \dots, x_p\} \subset K$  is an  $\epsilon$ -net for  $K$ . Let  $f$  be any Lipschitz function with domain including  $K$ .

**THEOREM 1.** *If  $\Phi: H \rightarrow H$  is any bounded linear transformation satisfying*

$$\Phi \mathbf{x}_i = \mathbf{f}(\mathbf{x}_i), \quad i = 1, \dots, p,$$

*then*

$$\sup_{x \in K} \|\varphi(x) - f(x)\| \leq (|\Phi| + |f|) \epsilon', \quad (11)$$

*where  $\varphi(x) = (\Phi\tau)(x)$  and  $\epsilon'$  can be made arbitrarily small.*

In short, if bounded linear  $\Phi$  interpolates the set  $\{(\mathbf{x}_i, \mathbf{f}(\mathbf{x}_i)); i = 1, \dots, p\}$  then  $\varphi$  interpolates the set  $\{(x_i, f(x_i)); i = 1, \dots, p\}$  and moreover approximates  $f$  on  $K$ . We note that  $\varphi$  is *not* linear.

*Proof.* In view of Lemma 2,  $\{\mathbf{x}_1, \dots, \mathbf{x}_p\}$  is an  $\epsilon'$ -net for  $\mathbf{K}$ , when  $\epsilon < \frac{1}{2}$  in fact  $\epsilon' < 2\epsilon$ . For arbitrary  $\mathbf{x} \in \mathbf{K}$  there exists an  $\mathbf{x}_j$  from the net satisfying  $\|\mathbf{x} - \mathbf{x}_j\| < \epsilon'$ . Now

$$\Phi \mathbf{x} - \mathbf{f}(\mathbf{x}) = [\Phi \mathbf{x} - \Phi \mathbf{x}_j] + [f(\mathbf{x}_j) - \mathbf{f}(\mathbf{x})],$$

where  $\Phi(\mathbf{x}_j) = \mathbf{f}(\mathbf{x}_j)$  has been used. Standard norm inequalities and the Lipschitz property of  $\Phi$  and  $\mathbf{f}$  establishes

$$\sup_{\mathbf{x} \in \mathbf{K}} \|\Phi(\mathbf{x}) - \mathbf{f}(\mathbf{x})\| \leq (|\Phi| + |\mathbf{f}|) \epsilon'.$$

Using the fact that  $\tau$  carries  $K$  one to one and onto  $\mathbf{K}$  and  $\mathbf{f}(\mathbf{x}) = f(x)$ ,  $\Phi(\mathbf{x}) = \varphi(x)$ , and  $|\mathbf{f}| \leq |f|$ , Eq. (11) follows.

**EXAMPLE.** To illustrate Theorem 1 we consider the causal Lagrange interpolation developed in [6]. For this we take  $v = [0, \infty]$ ,  $H = L_2(v)$ . Let  $K \subset H$  be compact and for  $\epsilon > 0$  let  $\{u_1, \dots, u_p\}$  be a set satisfying

- (i)  $\{u_1, \dots, u_p\}$  is an  $\epsilon$ -net for  $K$ ,
- (ii)  $\{P^t u_1, \dots, P^t u_p\}$  is a distinct set for  $t > 0$ .

The  $n^2$  functions,  $m_{ij}$ , are defined by

$$\begin{aligned} m_{ij}(u; t) &= \frac{\langle P^t(u - u_j)(u_i - u_j) \rangle}{\|P^t(u_i - u_j)\|^2}, & u \in H, \quad t \in v - 0, \\ &= 0, & t = 0. \end{aligned} \quad (12)$$

Each  $m_{ij}$  is scalar-valued continuous in  $t$ . The  $m_{ij}$  functions are intermediary to the functions

$$M_i(u; t) = \prod_{j \neq i} m_{ij}(u; t), \quad i = 1, \dots, p. \quad (13)$$

The relevant properties of the  $M_i$  include

- (i)  $P^t u = P^t v \Rightarrow M_i(u; \beta) = M_i(v, \beta), \beta \leq t$ ,
- (ii)  $M_i(u_j; t) = \delta_{ij}, t \in v$ ,
- (iii)  $M_i(u, \cdot)$  is continuous on  $v - 0$ .

Let  $f$  be an arbitrary causal, uniformly continuous function with domain containing  $K \cup \{u_1, \dots, u_p\}$ . Let  $y_i = f(u_i), i = 1, \dots, p$ , and denote by  $dP$  the operator-valued measure associated with the resolution of the identity given Eq. (3). From [6] we have the theorem:

**THEOREM 2** [6]. *The causal Lagrange interpolation of the well-posed problem given by*

$$\varphi(u) = \sum_{i=1}^p \int dP(s) y_i M_i(u; s), \quad u \in H,$$

*satisfies the causal interpolation problem.*

Now Theorem 1 states that if a linear  $\Phi: \mathbf{H} \rightarrow H$  exists such that  $\varphi = \Phi\tau$ , then  $\varphi$  approximates  $f$  on  $K$ . The existence of  $\Phi$  follows from inspection. Indeed with  $p = 3$  it follows from Eqs. (12) and (13) that  $M_1$ , for instance, expands to the form

$$\begin{aligned} M_1(u; t) = & K(t) \{ \langle P^t u, (u_1 - u_2) \rangle \langle (u_1 - u_3), P^t u \rangle \\ & - \langle P^t u, (u_1 - u_2) \rangle \langle P^t u_3, (u_1 - u_3) \rangle \\ & + \langle P^t u, (u_1 - u_3) \rangle \langle P^t u_2, (u_1 - u_2) \rangle \\ & + \langle P^t u_2, (u_1 - u_2) \rangle \langle P^t u_3, (u_1 - u_3) \rangle \}, \end{aligned}$$

where

$$K(t) = \|P^t(u_1 - u_2)\|^{-2} \|P^t(u_1 - u_3)\|^{-2}.$$

Using the concrete form of  $H = L_2$

$$\begin{aligned} & \int dP(s) y_1 M_1(u; s) \\ &= \int_0^t \int_0^t \omega_2(t, s_1, s_2) u(s_1) u(s_2) ds_1 ds_2 + \int_0^t \omega_1(t, s) u(s) ds + \omega_0(t), \quad t > 0, \end{aligned} \quad (14)$$

where

$$\begin{aligned}\omega_2(t, s_1, s_2) &= y_1(t) K(t) [(u_1(s_1) - u_2(s_1)) (u_1(s_2) - u_2(s_2))], \\ \omega_1(t, s) &= y_1(t) K(t) \left\{ (u_1(s) - u_2(s)) \int_0^t u_3(\beta) [u_1(\beta) - u_3(\beta)] d\beta \right. \\ &\quad \left. + (u_1(s) - u_3(s)) \int_0^t u_2(\beta) [u_1(\beta) - u_2(\beta)] d\beta \right\}, \\ \omega_0(t) &= y_1(t) K(t) \left\{ \int_0^t u_2(\beta) [u_1(\beta) - u_2(\beta)] d\beta \right. \\ &\quad \left. + \int_0^t u_3(\beta) [u_1(\beta) - u_3(\beta)] d\beta \right\}.\end{aligned}$$

Since all  $M_i$  are similar in form, and since Eq. (14) is obviously a linear map on  $\mathcal{H}$  the demonstration of  $\Phi$  is complete.

#### 4. INTERPOLATORS THAT OPTIMALLY APPROXIMATE

Returning now to Theorem 1 we consider the merits of the various possible linear  $\Phi$ 's satisfying the theorem. One property which is obviously desirable is that  $|\Phi|$  be small. A second desirable property is that the ratio  $\epsilon'/\epsilon$  be minimized. In view of Eq. (10) this ratio is minimized when  $n$  is the smallest integer for which a  $\Phi$  can be constructed. The above observations point toward the following considerations.

Let  $\mathfrak{J}$  be a finite index set. Using the multipower maps

$$(P_j u)(\cdot) = \int_v \cdots \int \varphi_j(\cdot, \alpha_1 \cdots \alpha_j) u(\alpha_1) \cdots u(\alpha_j), \quad j = 0, 1, \dots,$$

we form the polymonic operators

$$\varphi(\mathfrak{J}; u)(\cdot) = \sum_{j \in \mathfrak{J}} (P_j u)(\cdot), \quad u \in H. \quad (15)$$

The kernels of  $\varphi_j$  are assumed to be Hilbert-Schmidt type and we introduce the norms

$$\begin{aligned}|P_j|^2 &= \int_v \cdots \int |\varphi_j(t, \alpha_1, \dots, \alpha_j)|^2 dt d\alpha_1 \cdots d\alpha_j, \\ |\varphi(\mathfrak{J}; \cdot)|^2 &= \sum_{j \in \mathfrak{J}} |P_j|^2.\end{aligned} \quad (16)$$



**PROBLEM.** For each  $\mathbb{J}$  determine whether a  $\varphi(\mathbb{J}; \cdot)$  exists satisfying Eq. (1). If such a  $\varphi(\mathbb{J}; \cdot)$  exists determine the one with minimum norm.

We note that Eq. (15) computes also a linear map  $\Phi(\mathbb{J})$  with domain  $\mathbf{H}(\mathbb{J})$ . Here  $\mathbf{H}(\mathbb{J})$  denotes the vectorization of  $H$  with nonzero entries in the  $j$ th component only if  $j \in \mathbb{J}$ . It is not difficult to show that  $\|\Phi(\mathbb{J})\| = \|\varphi(\mathbb{J}; \cdot)\|$ .

For given  $\mathbb{J}$  and  $\{u_1, \dots, u_p\}$  let

$$\mu_{ij}(\mathbb{J}; t) = \sum_{k \in \mathbb{J}} \langle u_i, P^t u_j \rangle^k, \quad i, j = 1, \dots, p, \quad t \in \nu.$$

The matrix  $V(\mathbb{J}; t)$  is formed using  $\mu_{ij}$  as the row-column entries. When  $\mathbb{J} = \{1\}$  we see that  $V(\mathbb{J}; t)$  is the Gramian matrix of the set  $\{P^t u_1, \dots, P^t u_p\}$ .

**Result 1** (see [10]).

- (a) If  $\mathbb{J}' \subset \mathbb{J}$  then  $\text{rank } V(\mathbb{J}'; t) \leq \text{rank } V(\mathbb{J}; t)$ , all  $t \in \nu$ .
- (b) If  $t' \leq t$  then  $\text{rank } V(\mathbb{J}; t') \leq \text{rank } V(\mathbb{J}; t)$ , all  $\mathbb{J}$ .

A rather direct proof of Result 1 can be constructed along the following lines. It can be verified that  $V(\mathbb{J}; t)$  is the Gramian of the vectorized set  $\{\mathbf{P}^t \mathbf{u}_1, \dots, \mathbf{P}^t \mathbf{u}_p\} \subset \mathbf{H}(\mathbb{J})$ . Result 1(a) reflects the fact that allowing more nonzero components in the vectorized set will not decrease the linear independence. Result 1(b) reflects the fact that longer time histories also do not decrease linear independence.

For our second result we need the vector  $\mathbf{y}(t) = \text{col}(y_1(t), \dots, y_p(t))$ , where  $y_i = f(u_i)$ ,  $i = 1, \dots, p$ .

**Result 2** (see [10]). A causal solution to the polynomial interpolation over  $\mathbb{J}$  exists if and only if  $\mathbf{y}(t) \in \text{Range } V(\mathbb{J}; t)$ , a.e.  $t \in \nu$ .

One consequence of this result is that for the sets

$$\mathbb{J} = \{0\}, \dots, \{0, 1, \dots, m\}$$

a natural question is to find the smallest  $m$  such that  $V(\{0, \dots, m\}; t)$  is non-singular. The causal Lagrange interpolator examined earlier identifies sufficient conditions for  $m \leq p - 1$ . Examples can be constructed for which this limit is taken on; however, in most cases  $m$  is small relative to  $p$ .

**Result 3** (see [10]). Suppose that  $\det V(\mathbb{J}; t) \neq 0$ ,  $t > 0$ . Then the unique minimal norm interpolating  $\varphi$  has kernels given by

$$\varphi_j(t, \alpha_1, \dots, \alpha_j) = \Pi_j(t, \alpha_1, \dots, \alpha_j) V^{-1}(\mathbb{J}; t) \mathbf{y}(t),$$

where

$$\Pi_j(t, \alpha_1, \dots, \alpha_j) = \text{row}((P^t u_1)(\alpha_1) \cdots (P^t u_1)(\alpha_j), \dots, P^t(u_p)(\alpha_1) \cdots (P^t u_p)(\alpha_j)), \quad j \geq 1,$$

and

$$\Pi_0 = \text{row}(1, 1, \dots, 1).$$

The explicit form of Result 3 provides a direct computation of  $|\varphi|$ . Indeed it is easy to verify that

$$\sum_{j \in \mathbb{N}} \int \cdots \int |\varphi_j(t, \alpha_1, \dots, \alpha_j)|^2 d\alpha_1 \cdots d\alpha_j = \mathbf{y}^*(t) V^{-1}(\mathbb{N}; t) \mathbf{y}(t)$$

and hence

$$|\varphi|^2 = \int_v \mathbf{y}^*(t) V^{-1}(\mathbb{N}; t) \mathbf{y}(t) dt. \quad (17)$$

Since Eq. (15) identifies also a linear map  $\Phi(\mathbb{N})$  on  $\mathbf{H}(\mathbb{N})$  with  $|\varphi| = |\Phi|$  these results can be summarized as follows.

Let  $f$  be a causal function of  $H$  which is Lipschitz on  $K \cup \{u_1 \cdots u_p\} \subset H$ . For  $\epsilon > 0$  let  $\{u_1, \dots, u_p\}$  be an  $\epsilon$ -net for  $K$ . The index set  $\mathbb{N}$  has the property that  $V(\mathbb{N}; t)$  is invertible.

**THEOREM 3.** *The unique solution to the problem of Eq. (1) with constraints (2i), (2iii), (2iv), (2v), and  $\mathbb{N}$  specified is  $\varphi(\mathbb{N}; \cdot)$ . Moreover,*

$$\sup_{x \in K} \|f(x) - \varphi(\mathbb{N}; x)\| \leq \left[ \left( \int_v \mathbf{y}^*(t) V^{-1}(\mathbb{N}; t) \mathbf{y}(t) dt \right)^{1/2} + |f| \right] \epsilon',$$

where

$$(\epsilon')^2 = \sum_{\substack{j \in \mathbb{N} \\ j \neq 0}} \epsilon^{2j}.$$

**State realization.** It should be noted that the specification of the interpolator-approximator of this section is entirely independent of the state variable context. We show now that a state variable realization can also be obtained.

Suppose that  $z(t) = \varphi(\mathbb{N}; u)(t)$  and using Result 3 and Eqs. (15), (16) it follows that

$$z(t) = \mathbf{y}^*(t) \eta(t),$$

where

$$\eta(t) = V^{-1}(\mathbb{N}; t) \sum_{j \in \mathbb{N}} \int_0^t \cdots \int_0^t \Pi_j^*(\beta_1, \dots, \beta_j) u(\beta_1) \cdots u(\beta_j) d\beta_1 \cdots d\beta_j. \quad (18)$$

Using the explicit form of  $\Pi_j$  we see that

$$\begin{aligned} & \int_0^t \cdots \int_0^t \Pi_j^*(\beta_1, \dots, \beta_j) u(\beta_1) \cdots u(\beta_j) d\beta_1 \cdots d\beta_j \\ &= \text{col} \left( \cdots, \left[ \int_0^t u_\alpha(\beta) u(\beta) d\beta \right]^j, \cdots \right), \quad j = 0, 1, \dots, k. \end{aligned} \quad (19)$$

In the Appendix the computation of  $V^{-1}(\mathfrak{N}; t)$  is considered. The simplest case is when  $V(\mathfrak{N}; t)$  is invertible for all  $t \in \nu - 0$  and we restrict attention to this case here, as in the Appendix. The result we need from the Appendix is that there exist matrices  $F(t)$ ,  $\Phi(t)$ ,  $E(t)$  such that (Eq. (A3))

$$\frac{d}{dt} V^{-1}(\mathfrak{N}; t) = -V^{-1}(\mathfrak{N}; t) F^*(t) \Phi(t) F(t) V^{-1}(\mathfrak{N}; t), \quad (20)$$

where

$$\frac{d}{dt} \Phi(t) = E^*(t) \Phi(t) E(t). \quad (21)$$

For convenience we define the diagonal matrix

$$V(t) = \text{diag}[u_1(t), \dots, u_p(t)]$$

and the memoryless nonlinear function

$$\Psi[\lambda](t) = \sum_{(j+1) \in \mathfrak{N}} (j+1) V(t) \text{col}(\lambda_1^j, \lambda_2^j, \dots, \lambda_p^j), \quad \lambda \in R^p.$$

Using Eqs. (20), (21) and ordinary differentiation the following result can be verified.

**THEOREM 4.** Suppose  $z(t) = \varphi(\mathfrak{N}; u)(t)$  for the  $\varphi$  of Theorem 3. Let

$$\begin{aligned} \dot{\xi}(t) &= U(t) b u(t), \\ x(t) &= \Psi[\xi(t)], \end{aligned} \quad (22)$$

where  $b = \text{col}(1, 1, \dots, 1)$ . Then

$$\begin{aligned} \dot{\eta}(t) &= -V^{-1}(\mathfrak{N}; t) F^*(t) \Phi(t) F(t) \eta(t) + V^{-1}(\mathfrak{N}; t) x(t) u(t), \\ z(t) &= \mathbf{y}^*(t) \eta(t). \end{aligned} \quad (23)$$

The matrices  $\Phi(t)$ ,  $V^{-1}(\mathfrak{N}; t)$  are computed by Eqs. (20), (21).

It is noted that this realization is that of two linear systems coupled by a memoryless nonlinear function.

## 5. CONCLUSION

To place our mathematical results in a systems context we return to the discussion of Section 1. Consider, for example, a natural system such as the neuromuscular system of the human arm. In modeling such a system one can envision using selected input-output tests. The complicated internal interactions of such a system preclude any possibility of using conventional models.

The information in hand is likely to be of the following sort. If the neuromuscular system is preconditioned appropriately then the input-response pairs  $\{(u_i, y_i): i = 1, \dots, n\}$  are experimentally repeatable. The system is assumed to be causal and continuous but no other a priori information is available. A model is to be constructed which reproduces the existing experimental data. It is helpful if the modeling process is repeatable and if the resultant model has a minimal sensitivity to the experimental process.

Our results provide an effective response to such problems. Theorem 1 shows that interpolators which can be linearly realized on a vectorized space have an approximation property. Specifically the construction of Section 4 and [10] interpolates the existing data. For any  $\epsilon > 0$  and any input set such that  $\{u_i\}$  is an  $\epsilon$ -net the construction also approximates in an optimal way the neuromuscular system. This latter property provides the minimal sensitivity to experimental variation.

APPENDIX: COMPUTING  $V^{-1}(\mathbb{I}; t)$ 

In this Appendix we show that  $V^{-1}(\mathbb{I}; t)$  can be computed by a matrix differential equation. For convenience consider the polynomial

$$p(x) = \sum_{k \in \mathbb{I}} x^k, \quad x \in R.$$

The entries of the matrix  $V(\mathbb{I}; t)$  can be written in the form

$$V(\mathbb{I}; t)_{ij} = p(\langle x_i, P^t x_j \rangle), \quad t \in v.$$

Using calculus we have

$$\frac{d}{dt} V(\mathbb{I}; t)_{ij} = p^1(\langle x_i, P^t x_j \rangle) x_i(t) x_j(t), \quad t \in v, \quad (\text{A1})$$

where  $p^1(x)$  is the ordinary derivative of  $p$ .

With Eq. (A1) as a motivation let  $\max(\mathbb{I}) = k$  and let

$$p^j(x), \quad j = 0, 1, \dots, k,$$

denote the derivatives of  $p(x)$ , here  $p^0(x) = p(x)$  and  $p^k(x) = k!$ . Using these polynomials we define the  $k + 1$  matrices  $\Phi^0, \dots, \Phi^k$ , where  $\Phi^0 = V$  and

$$[\Phi^\alpha(t)]_{ij} = p^\alpha(\langle x_i, P^t x_j \rangle), \quad i, j = 1, \dots, n.$$

The diagonal matrix  $X(t) = \text{diag}(x_1(t), \dots, x_n(t))$  is needed as is the  $n(k + 1) \times n(k + 1)$  block subdiagonal matrix

$$E = \begin{bmatrix} 0 & & 0 \\ X & 0 & 0 \\ 0 & & X & 0 \end{bmatrix}.$$

By direct inspection it can be verified that

*Result.*  $(d/dt)\Phi = E^*\Phi E$ , where  $\Phi$  is the  $n(k + 1) \times n(k + 1)$  block diagonal matrix

$$\Phi = \text{diag}(\Phi^0, \dots, \Phi^k).$$

On a component basis

$$\begin{aligned} \frac{d}{dt} \Phi^0(t) &= X(t) \Phi^1(t) X(t), \\ \frac{d}{dt} \Phi^j(t) &= X(t) \Phi^{j+1}(t) X(t), \quad j = 1, \dots, k-1, \\ \frac{d}{dt} \Phi^k(t) &= 0. \end{aligned} \tag{A2}$$

Using  $V = \Phi^0$  and  $(d/dt)\{V^{-1}\} = -V^{-1}\Phi^0 V^{-1}$  we see that

$$\begin{aligned} \frac{d}{dt} \{V^{-1}(\mathbb{I}; t)\} &= -V^{-1}(\mathbb{I}; t) F(t)^* \Phi(t) F(t) V^{-1}(t), \\ \frac{d}{dt} \{\Phi(t)\} &= E^*(t) \Phi(t) E(t) \end{aligned} \tag{A3}$$

where  $F(t)$  is the  $n(k + 1) \times n$  block column matrix

$$F(t) = \text{col}(0, X(t), 0, 0, \dots, 0).$$

#### REFERENCES

1. R. K. MEHRA, Optimal input signals for Parameter estimation in systems-survey and new results, *IEEE Trans. Automatic Control* **AC-19**, No. 6 (Dec. 1974).
2. W. A. PORTER, An overview of polynomial system theory, *IEEE Proc. (Special Issue on System Theory)*, Jan. 1976.
3. G. PALM AND T. POGGIO, The Volterra representation and the Wiener expansion: Validity and pitfalls, *SIAM J. Appl. Math.* **33**, No. 2 (Sept. 1977).

4. W. A. PORTER, Partial system inverses for adaptation, parameter Estimation and sensitivity reduction, *IEEE Circuit System Trans.* (Special Issue: Math. Foundation of *System Theory*, 1978).
5. W. A. PORTER, The interrelation between observers and system sensitivity, *IEEE Trans. Automatic Control* **AC-22**, No. 1 (1977).
6. W. A. PORTER, Data interpolation, causality structure, and system identification, *Inform. Contr.* **29**, No. 3 (1975).
7. W. A. PORTER, Causal realization from input-output pairs, *SIAM J. Control* **15**, No. 1 (Jan. 1977).
8. W. A. PORTER, T. M. CLARK, AND R. M. DeSANTIS, Causality structure and the Weierstrass theorem, *J. Math. Anal. Appl.* **52** (1975).
9. W. A. PORTER, Approximation by Bernstein systems, *Math. Systems Theory* **11** (1978), 259-274.
10. W. A. PORTER, Synthesis of polynomic systems, *SIAM J. Appl. Math.*, in press.